

# Dzianis Prakapenka

## SUMMARY

---

I am an experienced developer with over 13 years of software development experience. I have a strong background in developing and optimizing computing pipelines, tools, and methods for working with large datasets. I am well versed in Python, C++, Fortran, and R (specifically for high performance scientific computing tasks).

## EDUCATION

---

**University of Minnesota** Minneapolis & Saint Paul, MN  
*Ph.D. in Bioinformatics and Computational Biology* Aug 2022

- **Thesis:** Haplotype Computing Tools for Genomic Prediction and Estimation and Detection of Epistasis Effects in Holstein Cattle

**University of Minnesota** Minneapolis & Saint Paul, MN  
*M.S. in Bioinformatics and Computational Biology* March 2019

- **Concentrations and coursework:** Mixed Model Methods for Genetic Analysis, Computation and Biology, Genetics, Statistical Genetics and Genomics, Computational Techniques in Genomics, Matrix Theory

**University of Chicago** Chicago, IL  
*B.A. in Economics* Dec 2013

- **Concentrations and coursework:** Econometrics, Statistical Models and Methods, Analysis in R

## COMPUTING SKILLS

---

**Languages and Software:** Python, C/C++, Fortran MKL | LAPACK, ScaLAPACK, OpenMP/MPI | R, Matlab, Mathematica, Stata | Java, Visual Basic, Basic, IDL | Linux, OS X, Windows, Chrome OS | Bash, Slurm Workload Manager, Portable Batch System

## EXPERIENCE

---

**University of Minnesota** Minneapolis & Saint Paul, MN  
*Postdoctoral Associate in Quantitative Genetics* Sept 2022 – Present

- Integrating complex epistasis effects in GVCHAPmpi genomic prediction software.
- Evaluating and integrating GPU acceleration to improve efficiency of genomic prediction tools.
- Conducting comprehensive, large scale haplotype analysis of all traits in USDA Holstein database

**University of Minnesota** Minneapolis & Saint Paul, MN  
*Research Assistant in Quantitative Genetics* Sept 2014 – Sept 2022

- Authored and published GVCHAP - a computing pipeline written in C++, Python, and R to integrate haplotypes with SNP markers for genomic prediction, estimation of variance components, and heritability.
- Developed and optimized computing tools and methods for quantitative analysis using SNPs and haplotypes, conducted multiple genomic prediction and variance component estimation analyses, and performed Genome-wide Association Studies (GWAS) and subsequent analyses for publications.
- Published a total of 13 peer-reviewed journal articles (3 first author) and conducted data analysis to secure funding from the NIH, NIFA, and USDA in collaboration with researchers from multiple fields on projects involving genomic selection and gene discovery in dairy cattle, swine, wheat, panda and human data.

**Research Computing Center at the University of Chicago** Chicago, IL  
*Research Assistant in Scientific Computing* Nov 2012 – Aug 2014

- Optimized and parallelized computational tasks to run on high-performance computing clusters.
- Redesigned applications to process intensive, time-consuming calculations faster with minimal memory usage.
- Collaborated with researchers in various disciplines to analyze and compute large data efficiently.

**Spallation Neutron Source at Oak Ridge National Laboratory** Oak Ridge, TN  
*Physics Programming Intern* June 2008 – Dec 2010

- Designed application to calculate and graphically display experimental data for interpretation.

- Programmed software to filter and catalog experimental data for users of neutron spectrometers.
- Developed tools to run spectrometry in collaboration with software engineers and researchers.

## Packer Engineering

Engineering Intern

Naperville, IL

Summer 2007

- Designed experiments, computer models, and simulations based on given scenarios.
- Conducted destructive and non-destructive systems testing and fracture/failure analysis of mechanical systems.

## University of Chicago

Laboratory Assistant/Programmer

Chicago, IL

Summer 2005

- Developed GUI and backend software for on-line synchrotron equipment (precise sample positioning, optical path correction.)
- Supported experiments for high-pressure X-ray analysis at the synchrotron.

## PEER-REVIEWED JOURNAL ARTICLES

---

- **Prakapenka D**, Liang Z, Da Y (2023) Genome-Wide Association Study of Age at First Calving in U.S. Holstein Cows. *Int. J. Mol. Sci.* 2023, 24, 7109. doi: 10.3390/ijms24087109
- **Prakapenka D**, Wang C, Liang Z, Bian C, Tan C and Da Y (2020) GVCHAP: A Computing Pipeline for Genomic Prediction and Variance Component Estimation Using Haplotypes and SNP Markers. *Front. Genet.* 11:282. doi: 10.3389/fgene.2020.00282
- **Prakapenka D**, Liang Z, Jiang J, Ma L, Da Y. (2021). A Large-Scale Genome-Wide Association Study of Epistasis Effects of Production Traits and Daughter Pregnancy Rate in U.S. Holstein Cattle. *Genes.* 2021;12(7). doi:10.3390/genes12071089
- Liang Z, **Prakapenka D**, Parker Gaddis KL, VandeHaar MJ, Weigel KA, Tempelman RJ, Koltjes JE, Santos JEP, Heather White, Francisco Peñagaricano, Ransom L Baldwin, Da, Y. (2022) Impact of epistasis effects on the accuracy of predicting phenotypic values of residual feed intake in U.S. Holstein cows. *Front. Genet.* doi: 10.3389/fgene.2022.1017490
- Da Y, Liang Z, **Prakapenka D**. (2022). Multifactorial Methods Integrating Haplotype and Epistasis Effects for Genomic Estimation and Prediction of Quantitative Traits. *Front. Genet.* doi: 10.3389/fgene.2022.922369
- Bian C, **Prakapenka D**, Tan C, Yang R, Zhu D, Guo X, Liu D, Cai G, Li Y, Liang Z, Wu Z, Da Y, Hu X. Haplotype genomic prediction of phenotypic values based on chromosome distance and gene boundaries using low-coverage sequencing in Duroc pigs. *Genet Sel Evol* 53, 78 (2021). doi: 10.1186/s12711-021-00661-y
- Liang, Z., C. Tan, **D. Prakapenka**, L. Ma, and Y. Da. 2020. Haplotype Analysis of Genomic Prediction Using Structural and Functional Genomic Information for Seven Human Phenotypes. *Frontiers in Genetics* 11(1461).
- Sallam, A. H., E. Conley, **D. Prakapenka**, Y. Da, and J. A. Anderson. 2020. Improving prediction accuracy using multi-allelic haplotype prediction and training population optimization in wheat. *G3: Genes, Genomes, Genetics* 10(7):2265-2273.
- Ma, L., T. S. Sonstegard, J. B. Cole, C. P. VanTassell, G. R. Wiggans, B. A. Crooker, C. Tan, **D. Prakapenka**, G. E. Liu, and Y. Da. 2019. Genome changes due to artificial selection in U.S. Holstein cattle. *BMC genomics* 20(1):128.
- Jiang J, Ma L, **Prakapenka D**, VanRaden PM, Cole JB, Da Y. A large-scale genome-wide association study in U.S. Holstein cattle. *Frontiers in Genetics.* 2019;10(412).
- Tan, C., Z. Wu, J. Ren, Z. Huang, D. Liu, X. He, **D. Prakapenka**, R. Zhang, N. Li, Y. Da, and X. Hu. 2017. Genome-wide association study and accuracy of genomic prediction for teat number in Duroc pigs using genotyping-by-sequencing. *Genetics Selection Evolution* 49:35.
- Garbe, J.R., **D. Prakapenka**, C. Tan, and Y. Da. 2016. Genomic inbreeding and relatedness in wild panda populations. *PLoS ONE* 11(8): e0160496.
- Wang, C., **D. Prakapenka**, S. Wang, S. Pulugurta, H.B. Runesha, and Y. Da. 2014. GVCBLUP: A computer package for genomic prediction and variance component estimation of additive and dominance effects. *BMC Bioinformatics* 15:270